

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/830,837A
Source: IFW/6
Date Processed by STIC: 11/10/04

ENTERED



IFW16

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,837A

DATE: 11/10/2004

TIME: 12:46:06

Input Set : A:\480848.90026.txt

Output Set: N:\CRF4\11102004\I830837A.raw

3 <110> APPLICANT: Institut de Recherches Cliniques de Montreal
 4 DAVIGNON, Jean
 5 LAAKSONEN, Reijo
 6 MARCINKIEWICZ, Mieczyslaw
 7 CHRETIEN, Michel
 8 SEIDAH, Nabil
 10 <120> TITLE OF INVENTION: MAMMALIAN SUBTILISIN/KEXIN ISOZYME SKI-1: A PROPROTEIN
 CONVERTASE

11 WITH A UNIQUE CLEAVAGE SPECIFICITY

13 <130> FILE REFERENCE: IRCM

C--> 15 <140> CURRENT APPLICATION NUMBER: US/09/830,837A

C--> 16 <141> CURRENT FILING DATE: 2001-10-18

18 <150> PRIOR APPLICATION NUMBER: CA 2,249,648

19 <151> PRIOR FILING DATE: 1998-11-04

21 <160> NUMBER OF SEQ ID NOS: 78

23 <170> SOFTWARE: PatentIn version 3.2

25 <210> SEQ ID NO: 1

26 <211> LENGTH: 3895

27 <212> TYPE: DNA

28 <213> ORGANISM: Rattus sp.

31 <220> FEATURE:

32 <221> NAME/KEY: CDS

33 <222> LOCATION: (418)..(3573)

35 <400> SEQUENCE: 1

36	gcgagtaaac	atccccgaa	tggatacccg	aggcgtgttc	gcggcgagg	ccccgtttc	60		
38	ccgggtccgc	cgatcccgag	cctgaggcga	cgcagatcg	ctcagagcgg	tggctggc	120		
40	tcctgctaga	tttgggtctg	tggtacaaat	ggagtttagg	actcagtgga	ctcgcccta	180		
42	atgagagaag	ccccctgtcc	aagatggaga	agaagcggag	aaagaaatga	aagccttt	240		
44	ttgggccaag	ctgtgggtga	ccatggact	gaggtttct	ttacgttgga	caagtctgta	300		
46	ggatggctga	tcagtaaggt	tgcagttt	agcgaaaaca	gaaatccact	tctgatcaag	360		
48	gaagagccta	gtgcaatttg	aatttatgca	attttatgac	catattcaact	taggacc	417		
50	atg aag	ctc gtc	aac atc	tgg ctt	ctt ctg	ctg gtg	465		
51	Met Lys	Leu Val	Asn Ile	Trp Leu	Leu Val	Val Leu	Cys		
52	1	5	10	15					
54	ggg aaa aag	cat ctg	ggt gac	agg ctg	ggg aag	aaa gct	513		
55	Gly Lys	Lys His	Leu Gly	Asp Arg	Leu Gly	Lys Ala	Phe Glu Lys		
56	20	25	30						
58	gcc cca	tgc ccc	agc tgt	tcc cac	ctg act	ttg aag	gtg gaa	ttc tcc	561
59	Ala Pro	Cys Pro	Ser Cys	Ser His	Leu Thr	Leu Lys	Val Glu	Phe Ser	
60	35	40	45						
62	tca act	gtg gtg	gaa tat	gaa tat	att gtg	gct ttc	aac gga	tac ttc	609
63	Ser Thr	Val Val	Glu Tyr	Glu Tyr	Ile Val	Ala Phe	Asn Gly	Tyr Phe	
64	50	55	60						
66	aca gcc	aaa gct	aga aac	tca ttt	att tca	agt gct	cta aaa	agc agt	657

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/830,837A

DATE: 11/10/2004
TIME: 12:46:06

Input Set : A:\480848.90026.txt
Output Set: N:\CRF4\11102004\I830837A.raw

67	Thr	Ala	Lys	Ala	Arg	Asn	Ser	Phe	Ile	Ser	Ser	Ala	Leu	Lys	Ser	Ser	
68	65				70					75					80		
70	gaa	gtg	gac	aac	tgg	aga	ata	ata	cct	cg	aac	aac	cca	tct	agt	gac	705
71	Glu	Val	Asp	Asn	Trp	Arg	Ile	Ile	Pro	Arg	Asn	Asn	Pro	Ser	Ser	Asp	
72					85				90				95				
74	tac	cct	agt	gat	ttt	gag	gtg	att	cag	ata	aaa	gag	aag	cag	aag	gcg	753
75	Tyr	Pro	Ser	Asp	Phe	Glu	Val	Ile	Gln	Ile	Lys	Glu	Lys	Gln	Lys	Ala	
76					100				105				110				
78	ggg	ctg	ctc	aca	ctt	gaa	gat	cac	cca	aac	atc	aag	cg	gtg	aca	ccc	801
79	Gly	Leu	Leu	Thr	Leu	Glu	Asp	His	Pro	Asn	Ile	Lys	Arg	Val	Thr	Pro	
80					115				120				125				
82	cag	cg	aaa	gtc	ttt	cgt	tcc	ctg	aag	ttt	gct	gaa	tcc	gac	ccc	att	849
83	Gln	Arg	Lys	Val	Phe	Arg	Ser	Leu	Lys	Phe	Ala	Glu	Ser	Asp	Pro	Ile	
84					130				135				140				
86	gtg	ccc	tgt	aat	gag	acc	cg	tgg	agc	cag	aag	tgg	cag	tca	tca	cgt	897
87	Val	Pro	Cys	Asn	Glu	Thr	Arg	Trp	Ser	Gln	Lys	Trp	Gln	Ser	Ser	Arg	
88					145				150				155			160	
90	ccc	ctg	aaa	aga	gcc	agt	ctc	tcc	ctg	ggc	tct	gga	ttc	tgg	cat	gca	945
91	Pro	Leu	Lys	Arg	Ala	Ser	Leu	Ser	Leu	Gly	Ser	Gly	Phe	Trp	His	Ala	
92					165				170				175				
94	aca	gga	agg	cat	tca	agt	cga	cga	ttg	ctg	aga	gcc	att	cct	cgc	cag	993
95	Thr	Gly	Arg	His	Ser	Ser	Arg	Arg	Leu	Leu	Arg	Ala	Ile	Pro	Arg	Gln	
96					180				185				190				
98	gtt	gcc	cag	aca	ttg	cag	gca	gat	gtg	ctt	tgg	cag	atg	gga	tac	aca	1041
99	Val	Ala	Gln	Thr	Leu	Gln	Ala	Asp	Val	Leu	Trp	Gln	Met	Gly	Tyr	Thr	
100					195				200				205				
102	ggt	gct	aat	gtc	agg	gtt	gcc	gtt	ttt	gat	act	ggg	ctc	agt	gag	aag	1089
103	Gly	Ala	Asn	Val	Arg	Val	Ala	Val	Phe	Asp	Thr	Gly	Leu	Ser	Glu	Lys	
104					210				215				220				
106	cat	cca	cat	ttc	aag	aat	gtg	aag	gaa	aga	acc	aac	tgg	acc	aat	gag	1137
107	His	Pro	His	Phe	Lys	Asn	Val	Lys	Glu	Arg	Thr	Asn	Trp	Thr	Asn	Glu	
108					225				230				235			240	
110	cgg	acc	ctg	gac	gat	ggg	ctg	ggc	cat	ggc	aca	tcc	gtt	gca	ggt	gtg	1185
111	Arg	Thr	Leu	Asp	Asp	Gly	Leu	Gly	His	Gly	Thr	Phe	Val	Ala	Gly	Val	
112					245				250				255				
114	att	gcc	agc	atg	aga	gag	tgc	caa	gga	ttt	gcc	cca	gat	gca	gag	ctg	1233
115	Ile	Ala	Ser	Met	Arg	Glu	Cys	Gln	Gly	Phe	Ala	Pro	Asp	Ala	Glu	Leu	
116					260				265				270				
118	cac	atc	ttc	agg	gtc	ttt	acc	aac	aat	cag	gtg	tct	tac	acg	tct	tgg	1281
119	His	Ile	Phe	Arg	Val	Phe	Thr	Asn	Asn	Gln	Val	Ser	Tyr	Thr	Ser	Trp	
120					275				280				285				
122	ttt	ttg	gat	gcc	ttc	aac	tat	gcc	atc	cta	aag	aag	atg	gac	gtt	ctg	1329
123	Phe	Leu	Asp	Ala	Phe	Asn	Tyr	Ala	Ile	Leu	Lys	Lys	Met	Asp	Val	Leu	
124					290				295				300				
126	aac	ctt	agc	atc	ggt	ggg	cct	gac	ttc	atg	gat	cac	ccc	ttt	gtt	gac	1377
127	Asn	Leu	Ser	Ile	Gly	Gly	Pro	Asp	Phe	Met	Asp	His	Pro	Phe	Val	Asp	
128					305				310				315			320	
130	aag	gta	tgg	gaa	tta	aca	gcg	aac	aat	gta	atc	atg	gtt	tct	gtt	att	1425
131	Lys	Val	Trp	Glu	Leu	Thr	Ala	Asn	Asn	Val	Ile	Met	Val	Ser	Ala	Ile	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,837A

DATE: 11/10/2004

TIME: 12:46:06

Input Set : A:\480848.90026.txt

Output Set: N:\CRF4\11102004\I830837A.raw

132	325	330	335	
134	ggc aat gat gga cct ctc tat ggc act ctg aat aac cct gct gat cag			1473
135	Gly Asn Asp Gly Pro Leu Tyr Gly Thr Leu Asn Asn Pro Ala Asp Gln			
136	340	345	350	
138	atg gat gtg att gga gtg ggt ggc att gac ttt gaa gac aac atc gcc			1521
139	Met Asp Val Ile Gly Val Gly Gly Ile Asp Phe Glu Asp Asn Ile Ala			
140	355	360	365	
142	cgc ttc tct tcc agg gga atg act acc tgg gaa cta ccg gga ggc tat			1569
143	Arg Phe Ser Ser Arg Gly Met Thr Thr Trp Glu Leu Pro Gly Gly Tyr			
144	370	375	380	
146	ggt cgt gtg aag cct gac att gtc acc tat ggt gct gga gtg cgg ggt			1617
147	Gly Arg Val Lys Pro Asp Ile Val Thr Tyr Gly Ala Gly Val Arg Gly			
148	385	390	395	400
150	tct ggt gtg aaa ggg ggc tgc cgt gca ctc tca ggg acc agt gtc gcc			1665
151	Ser Gly Val Lys Gly Gly Cys Arg Ala Leu Ser Gly Thr Ser Val Ala			
152	405	410	415	
154	tcc cca gtg gtt gct ggg gct gtc acc ttg tta gta agc aca gta cag			1713
155	Ser Pro Val Val Ala Gly Ala Val Thr Leu Leu Val Ser Thr Val Gln			
156	420	425	430	
158	aag cgg gag cta gtg aat cct gcc agt gtg aag caa gct ttg ata gca			1761
159	Lys Arg Glu Leu Val Asn Pro Ala Ser Val Lys Gln Ala Leu Ile Ala			
160	435	440	445	
162	tca gcc cgg aga ctt cct ggt gtc aac atg ttt gag caa ggc cat ggc			1809
163	Ser Ala Arg Arg Leu Pro Gly Val Asn Met Phe Glu Gln Gly His Gly			
164	450	455	460	
166	aag ttg gat cta ctg cga gcc tat cag atc ctc agc agc tat aaa ccg			1857
167	Lys Leu Asp Leu Leu Arg Ala Tyr Gln Ile Leu Ser Ser Tyr Lys Pro			
168	465	470	475	480
170	cag gcg agc ctg agt cct agc tac atc gac ctg act gag tgt ccc tac			1905
171	Gln Ala Ser Leu Ser Pro Ser Tyr Ile Asp Leu Thr Glu Cys Pro Tyr			
172	485	490	495	
174	atg tgg ccc tac tgc tcc cag ccc atc tac tat gga gga atg cca aca			1953
175	Met Trp Pro Tyr Cys Ser Gln Pro Ile Tyr Tyr Gly Met Pro Thr			
176	500	505	510	
178	att gtt aat gtc acc atc ctc aat ggc atg gga gtt aca gga aga att			2001
179	Ile Val Asn Val Thr Ile Leu Asn Gly Met Gly Val Thr Gly Arg Ile			
180	515	520	525	
182	gtg gat aag cct gag tgg cga ccc tat tta cca cag aat gga gac aac			2049
183	Val Asp Lys Pro Glu Trp Arg Pro Tyr Leu Pro Gln Asn Gly Asp Asn			
184	530	535	540	
186	att gaa gtg gcc ttc tcc tac tcc tca gtg ttg tgg cct tgg tca ggt			2097
187	Ile Glu Val Ala Phe Ser Tyr Ser Ser Val Leu Trp Pro Trp Ser Gly			
188	545	550	555	560
190	tac ctt gcc atc tcc att tct gtg acc aag aag gca gct tcc tgg gaa			2145
191	Tyr Leu Ala Ile Ser Ile Ser Val Thr Lys Lys Ala Ala Ser Trp Glu			
192	565	570	575	
194	ggc atc gcg cag ggc cac atc atg atc aca gtg gct tcc cca gca gag			2193
195	Gly Ile Ala Gln Gly His Ile Met Ile Thr Val Ala Ser Pro Ala Glu			
196	580	585	590	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,837A

DATE: 11/10/2004

TIME: 12:46:06

Input Set : A:\480848.90026.txt

Output Set: N:\CRF4\11102004\I830837A.raw

198 acg gaa tta aaa aat ggt gcc gag cat act tcc aca gtg aag ctg ccc	2241
199 Thr Glu Leu Lys Asn Gly Ala Glu His Thr Ser Thr Val Lys Leu Pro	
200 595 600 605	
202 atc aag gtg aag atc att ccc acc cct cct cggt agc aag aga gtc ctc	2289
203 Ile Lys Val Lys Ile Ile Pro Thr Pro Pro Arg Ser Lys Arg Val Leu	
204 610 615 620	
206 tgg gac cag tac cac aac ctc cgc tac cca ccc ggc tac ttc ccc agg	2337
207 Trp Asp Gln Tyr His Asn Leu Arg Tyr Pro Pro Gly Tyr Phe Pro Arg	
208 625 630 635 640	
210 gac aac ttg cgg atg aag aat gat cct tta gac tgg aat ggc gac cac	2385
211 Asp Asn Leu Arg Met Lys Asn Asp Pro Leu Asp Trp Asn Gly Asp His	
212 645 650 655	
214 gtc cac acc aac ttc agg gac atg tac cag cat ctg cgc agc atg ggc	2433
215 Val His Thr Asn Phe Arg Asp Met Tyr Gln His Leu Arg Ser Met Gly	
216 660 665 670	
218 tac ttt gtg gag gtg ctt ggt gcc cca ttc aca tgc ttt gac gcc acg	2481
219 Tyr Phe Val Glu Val Leu Gly Ala Pro Phe Thr Cys Phe Asp Ala Thr	
220 675 680 685	
222 cag tac ggc act ctg ctt atg gtg gac agt gag gaa gag tac ttc cct	2529
223 Gln Tyr Gly Thr Leu Leu Met Val Asp Ser Glu Glu Glu Tyr Phe Pro	
224 690 695 700	
226 gag gag att gct aag ctg agg agg gac gtg gac aat ggc ctt tcc ctt	2577
227 Glu Glu Ile Ala Lys Leu Arg Arg Asp Val Asp Asn Gly Leu Ser Leu	
228 705 710 715 720	
230 gtc gtc ttc agt gac tgg tac aac act tct gtt atg aga aaa gtg aag	2625
231 Val Val Phe Ser Asp Trp Tyr Asn Thr Ser Val Met Arg Lys Val Lys	
232 725 730 735	
234 ttt tac gat gaa aac aca agg cag tgg tgg atg cca gat act gga gga	2673
235 Phe Tyr Asp Glu Asn Thr Arg Gln Trp Trp Met Pro Asp Thr Gly Gly	
236 740 745 750	
238 gcc aac gtc cca gct cta aac gag ctg ctg tct gtt tgg aac atg ggg	2721
239 Ala Asn Val Pro Ala Leu Asn Glu Leu Leu Ser Val Trp Asn Met Gly	
240 755 760 765	
242 ttc agt gac ggc ctg tat gaa ggg gag ttt gcc ctg gca aac cac gac	2769
243 Phe Ser Asp Gly Leu Tyr Glu Gly Glu Phe Ala Leu Ala Asn His Asp	
244 770 775 780	
246 atg tac tat gca tcg ggg tgc agc att gcc agg ttt cca gaa gat ggt	2817
247 Met Tyr Tyr Ala Ser Gly Cys Ser Ile Ala Arg Phe Pro Glu Asp Gly	
248 785 790 795 800	
250 gtg gtg atc aca cag act ttc aag gac caa gga ttg gaa gtc tta aaa	2865
251 Val Val Ile Thr Gln Thr Phe Lys Asp Gln Gly Leu Glu Val Leu Lys	
252 805 810 815	
254 caa gag aca gca gtt gtc gac aat gtc ccc att ctg ggg cta tat cag	2913
255 Gln Glu Thr Ala Val Val Asp Asn Val Pro Ile Leu Gly Leu Tyr Gln	
256 820 825 830	
258 att cca gct gaa ggt gga ggc cgg att gtg ctg tat gga gac tcc aac	2961
259 Ile Pro Ala Glu Gly Gly Arg Ile Val Leu Tyr Gly Asp Ser Asn	
260 835 840 845	
262 tgc ttg gat gac agt cac aga cag aag gac tgc ttt tgg ctt ctg gat	3009

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,837A

DATE: 11/10/2004

TIME: 12:46:06

Input Set : A:\480848.90026.txt
Output Set: N:\CRF4\11102004\I830837A.raw

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 11/10/2004
PATENT APPLICATION: US/09/830,837A TIME: 12:46:07

Input Set : A:\480848.90026.txt
Output Set: N:\CRF4\11102004\I830837A.raw

lease Note:

se of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> o <223> fields of each sequence which presents at least one n or Xaa.

eq#:7; Xaa Pos. 2,3,4,5,6,7
eq#:8; Xaa Pos. 2,3,4,5,6,7
eq#:9; Xaa Pos. 2,3,4,5,6,7,8
eq#:10; Xaa Pos. 2,3,4,5,6,7,8
eq#:11; Xaa Pos. 2,3,4,5,6,7,8,9
eq#:12; Xaa Pos. 2,3,4,5,6,7,8,9
eq#:14; Xaa Pos. 1,13
eq#:15; N Pos. 3,9,12,18,21
eq#:16; N Pos. 3,6,9,12,15,18,21,24,29
eq#:17; Xaa Pos. 5,6
eq#:18; Xaa Pos. 4,6,8,10
eq#:46; Xaa Pos. 1,13
eq#:47; Xaa Pos. 1,11
eq#:74; Xaa Pos. 1,13
eq#:75; Xaa Pos. 1,11
eq#:76; Xaa Pos. 1,14

nvalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

eq#: 77, 78

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/830,837A**

DATE: 11/10/2004
TIME: 12:46:07

Input Set : **A:\480848.90026.txt**
Output Set: **N:\CRF4\11102004\I830837A.raw**

L:15 M:270 C: Current Application Number differs, Replaced Current Application Number
L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1774 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
L:1794 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:1814 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0
L:1834 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0
L:1854 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:1874 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:1913 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0
L:1952 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0
L:2010 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0
L:2029 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0
L:2064 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:0
L:2402 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:0
L:2427 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:0
L:2741 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:0
L:2766 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75 after pos.:0
L:2791 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76 after pos.:0